Question Answering Systems: A Systematic Literature Review

Sarah Saad Alanazi[1]
Nazar Elfadil[2]
Department of Computer Science
Fahad Bin Sultan University
Tabuk, Saudi Arabia

Mutsam Jarajreh[3]
Computer Engineering Department
Fahad Bin Sultan University
Tabuk, Saudi Arabia

Saad Algarni[4]
Business Administration Department
Fahad Bin Sultan University
Tabuk, Saudi Arabia

Abstract—Question answering systems (QAS) are developed to answer questions presented in natural language by extracting the answer. The development of QAS is aimed at making the Web more suited to human use by eliminating the need to sift through a lot of search results manually to determine the correct answer to a question. Accordingly, the aim of this study was to provide an overview of the current state of QAS research. It also aimed at highlighting the key limitations and gaps in the existing body of knowledge relating to QAS. Furthermore, it intended to identify the most effective methods utilized in the design of QAS. The systematic review of literature research method was selected as the most appropriate methodology for studying the research topic. This method differs from the conventional literature review as it is more comprehensive and objective. Based on the findings, QAS is a highly active area of research, with scholars taking diverse approaches in the development of their systems. Some of the limitations observed in these studies encompass the focused nature of current QAS, weaknesses associated with models that are used as building blocks for QAS, the need for standard datasets and question formats hence limiting the applicability of the QAS in practical settings, and the failure of researchers to examine their QAS solutions comprehensively. The most effective methods for designing QAS include focusing on syntax and context, utilizing word encoding and knowledge systems, leveraging deep learning, and using elements such as machine learning and artificial intelligence. Going forward, modular designs ought to be encouraged to foster collaboration in the creation of QAS.

Keywords—*Question answering systems; syntax; knowledge systems; deep learning; machine learning; systematic literature review; artificial intelligence*

I. Introduction

Information retrieval has undergone tremendous transformation in the recent past. Today, modern information access systems enable us to retrieve documents that may be linked to the input we supply the system. However, in most cases, the user is left to extract the important information from the retrieved documents. For example, the question "Who has finished a marathon race in under two hours?" Should return the answer "Eliud Kipchoge". Instead, the user is supplied with a list of associated documents to explore and reveal the correct answer. Regardless of this limitation, Question Answering System (QAS) has been noted as an area with great potential in modern computing [3, 4, 5]. QAS allows access to information in a very natural way, that is, by asking questions and getting related responses in natural language [9, 10, 12, 14, 41, 70].

A. Current Rsearch Limitations

Because of the great potential and usefulness of QAS, it will definitely be a subject to major advancement in the forthcoming years. Nonetheless, there are significant challenges that will require urgent resolving if we are to attain full potential of QAS. One of these challenges is the existing asymmetry between natural language and machine language. The asymmetry has delayed the ability of question answering systems to understand natural language-based input from users and interpret it correctly for an accurate retrieval of responses [43, 44, 45, and 46]. The language asymmetry has been a compound of several factors such as classification, construction of correct questions, ambiguity resolution, deficiencies in semantic equivalence recognition, and poor identification of sequential association in complex queries [50, 51, 52, 53, 54, 55, 56, 57, 58, 59, and 71]. Besides, there has also been a lack of accurate validation mechanism to guarantee accurateness in the responses produced by QAS (689). The unresolved nature of most of these challenges in the QAS literature is the most significant limitation of QAS research.

B. Research Motivations

Despite the above limitation, among other challenges facing the development of perfect QAS, researchers have not given up on the quest to develop a more accurate QAS. Across all ages, human beings have always exhibited an acute thirst for information and a difference between this information and knowledge has always existed [17, 18, 19, 25, 32, 33, 35, 37, 38, 40]. Owing to this difference, consistent research has led to the maturity of modern information retrieval systems such as web searches, which allow users to access information related to their interests at their fingertips. QAS is a modern and specialized method of information access that seeks to bridge the gap between the information that users may have and the relevant knowledge. In a typical internet browsing session, an internet user is not interested in the relevant webpages that come up when making internet searches, rather, the interest is in the answers to the questions defining the searches. Bridging the gap between the questions and answers has been the main motivation of modern QAS research and the great potential of QAS makes it an exciting field to explore.

C. Problem Statement

The working mechanism of modern question answering systems can be broken down into three broad stages, namely question analysis, document analysis, and answer analysis. The question analysis stage entails the parsing and classification of questions, as well as the development of queries that can be interpreted by the machine. The second stage of document analysis involves the extraction of documents that are relevant to the questions as interpreted by the machine and identification of suitable answers. The third stage of answer analysis involves a further breakdown of the individual documents to extract candidate answers and rank them according to their relevancy to the question.

In all the three stages, a combination of techniques from AI, NLP, statistical processing, pattern identification and matching, and information retrieval and extraction is used [2, 74, and]. The majority of modern question answering systems incorporate most if not all the above techniques to deliver improved accuracy of results. In fact, the taxonomy of the modern QAS is derived from the techniques that forms the basis of the systems at different working stages. Such include, Linguistic-based QAS, Statistical-based QAS, and Pattern-based QAS approaches. Even the most sophisticated of these approaches has always faced the challenges of language asymmetry, among other challenges which have delayed the attainment of ultimate perfection in QASs. The desire to resolve the challenges has been a significant impetus for QAS research. A review of the literature on QAS is needed to understand the extent reached by current QAS research in resolving the outstanding challenges to the attainment of the perfect question answering system.

D. Research Objectives

This paper provides a systematic review of the modern question answering systems' literature. The areas of interest to the paper are the current state of QAS research and identification of the most significant gaps and limitations in the reviewed studies. The three objectives of this research are broken down into three research questions.

E. Research Questions

RQ1: What is the current state of QAS research?

RQ2: Which are the most significant gaps and limitations in the reviewed studies?

RQ3: What are the most effective techniques used in designing QAS?

F. Summary

The remainder of the paper is organized as follows; following in the next section is a review of the recent literature on QAS then a methodology. The research methodology section emphasizes on planning, the review process, and reporting the results of systematic review. Lastly, the paper offers a conclusion relevant to the three research questions and the reviewed literature.

II. RELATED WORKS

This section of the paper provides a review of modern studies on QAS. However, it is important to appreciate that the development of the modern question answering systems has not been an event, but a process with a rich history. Modern studies have built on the findings of older studies to better the perfection of modern information retrieval systems. Advanced research on QAS extends back four decades ago and has grown parallel to the whole natural language processing field. Over the last four decades, hundreds of question answering systems have been developed following tremendous research efforts.

QASs are information retrieval-based tasks that process questions posed in natural language by using pre-organized databases or a large corpus of documents published in natural language. In another way, QAS accepts questions in a natural language and returns a collection of related responses in natural language. The demand for systems with this capability has been exponential owing to the growing quest for precision in the wake of increasing data and information. This has transformed it into a field with growing academic research interest from around the world.

Like any other technical field in the modern day computing, there are key terms that are related to QAS. Defining these terms will help us understand the concepts used in QAS. One of the key terms in the QAS literature is "Question Phrase", which is the section of the question that contains the search items (636). Another term is the "Question Type" which is identifies the kind of question given its purpose (636). In the QAS literature, the "Answer Type" means the classification of items that the question is seeking (636). "Question Focus" is the property related to the items of the sought by the question and "Candidate Passage" are the items identified by the search system as relevant to the search question. A candidate passage can be anything from a document or sentence in natural language that is retrieved by the search system. A "Candidate Answer" is response ranked as among the most suitable answer to the search question.

QAS literature divides the working mechanism of question answering systems into three broad modules, namely, question processing, document processing, and answer processing. As noted earlier, the question processing stage entails the parsing and classification of questions, as well as the development of queries that can be interpreted by the machine. The goal of the first stage is to identify the type, which defines the focus of the question. The focus of the question is identified by classifying as a "What", "Why", "Who", "Where "or "How" question so that the expected answer can be determined (141). This is important in bettering answer detection, which ultimately leads to acceptable accuracy of the returned answers.

The role of the document processing stage is to select a collection document that are related to the question posed by the user. The document processing stage also involve extracting few paragraphs from the selected documents that conform to the focus of the question. In modern QAS, this stage generates a dataset or a neural model which acts as the pseudocode for the process of answer extraction [13, 76, 77, and 78]. The data that is retrieved in the second stage is organized with preference inclined to those that are highly relevant to the question.

The third stage of answer processing involves a further breakdown of the individual documents to extract candidate

answers and rank them according to their relevancy to the question (5). This is often the most challenging task of the three. It involves further analysis of the document analysis stage to select the most suitable answer to the question. The complexity emanates from the need to make the answer as simple as possible even when it requires combining of information from different neural models.

III. RESEARCH METHOD

This systematic review followed the guidelines provided eight steps, amongst which the most significant are the purpose for reviewing the literature, searching the literature, screening the literature, quality evaluation, and data abstraction. The flowing section outlines the stages of the systematic literature review conducted in completing the paper.

A. Planning the Review

The systematic review of literature started with the creation of an elaborate plan. The key components of the plan included identifying the resources required and defining the timeframes for the completion of the process. In addition to identifying the various scholarly databases, other components of the plan entailed timelines for deriving research questions from the topic, creating the search strategy, determining the search terms and strings, implementing the search strategy, selecting the most relevant studies, reviewing those studies, and writing the research paper. The detailed phases are provided in the following sections.

B. Specifying the Research Questions (RQS)

Over the years, we have seen an exponential expansion of digital databases that have increasingly pushed for sophisticated tools of information retrieval. With the data at hand, the challenge has always been to develop efficient techniques of consuming the data, which involves using the information we have to extract knowledge from the digital information. One of such techniques is the question answering system which allows human users to interact with computers in the most natural way as they seek answers to their questions from large corpus of unstructured data. In this review, we create three questions, as noted under section 1, subsection 4 to guide this systematic review of literature, which seeks to understand the current state of QAS research and identification of the most significant gaps and limitations in the reviewed studies.

RQ1: What is the current state of QAS research?

RQ2: Which are the most significant gaps and limitations in the reviewed studies?

RQ3: What are the most effective techniques (Method) used in designing QAS?

C. Defining Search Strategy

1) Data retrieval: The first step of the review was to index the journals and papers, including conference proceedings written and published in English. A date filter of the year of publication was limited to between 2015 and 2020. The journals and papers, including conference proceedings were pulled from five digital libraries, ACM Digital Library,

IEEE Xplore, Science Direct – Elsevier, Springer Link, and Wiley. This was achieved using a conceptual research string containing the keywords in the research questions.

2) Screening of the journals and papers: The choice of the papers included in this review were determined using an inclusion and exclusion criteria. The choice of the papers included in this review were determined using an inclusion and exclusion criteria to ensure that the study was explicit about the journals and papers included in the research. Only papers that met the criteria were included in the review. Table I provides more details about the inclusion and exclusion criteria used in recruiting reviewed literature.

TABLE I.   INCLUSION AND EXCLUSION CRITERIA DEFINED FOR SCREENING

Inclusion Criteria	Exclusion Criteria
Only papers written and published in the English language	Non-English academic works
Academic research work published in conferences and journals	Duplicate papers existing in separate libraries
Question related to QAS, particularly those touching on the current state of QAS research and those with the potential to reveal the most significant gaps and limitations in the reviewed studies	Books, thesis, editorials among others that do not constitute published academic research
QAS studies published after January 1, 2018	Academic works published before January 1, 2018

D. Defining Data Sources(Eligibility of the Journals and Papers)

To ensure that only relevant papers were included in the review. The scheme involved answering quality assessment questions with a yes = 1.5, partial = 0.5, a no = 0 depending on a preliminary analysis of the individual papers. The papers that had been preselected but had a score of less than 0.5 were excluded from the review which resulted in a sample 130 papers. Purposive exclusion of 50 papers was done to remain with 80 papers.

E. Defining Search Keywords

The initial search words were derived from the three research questions. Additional search words were determined based on the results of the initial search. Table II highlights the main search words and offers an explanation of each one of them.

TABLE II.   SEARCH KEYWORDS

QAS	Question answering systems
Syntax Knowledge systems Deep learning	Arrangement of phrases and words Collection of knowledge presented using some formal representation Artificial intelligence function that utilizes multiple layers to extract features from raw input.
Machine learning	Subset of artificial intelligence that entails utilizing statistical methods to enable machines learn automatically without explicit programming.
Artificial intelligence	Programming compiuters to mimic the behavior and thought of human beings.

Several search strings were developed using the search words shown in Table II and combined with Boolean operators. The utilization of Boolean operators, primarily AND and OR, was important in identifying the most appropriate studies.

F. Conducting Review Process

The process entailed identifying records, screening them, determining their eligibility, and listing the included studies in accordance with PRISMA. Fig. 1 summarizes the search protocol.
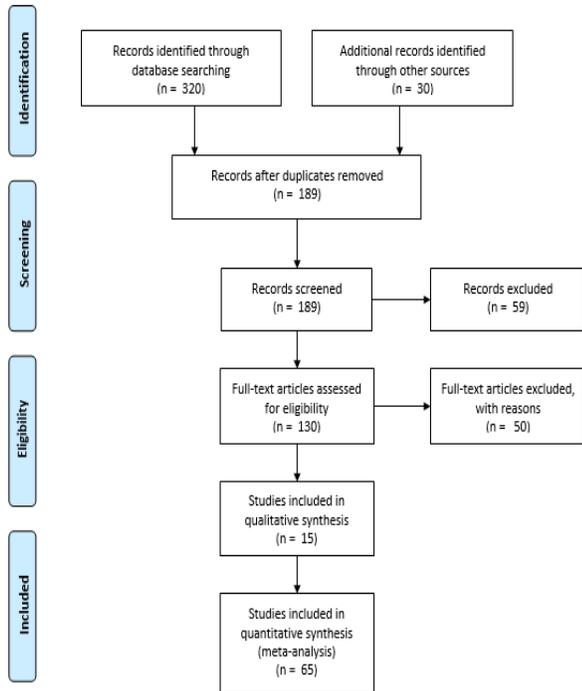
Fig. 1.   Search Strategy.

G. Selection of Study

The articles chosen for this study were determined using two-level inclusion and exclusion criteria. The identification process produced a total of 350 articles, with 320 being found through database searching whereas 30 were identified using other sources. After the exclusion of duplicates, 189 studies remained. The screening process resulted in the elimination of 59 studies. The resultant 130 studies were assessed for eligibility and 50 were excluded with reasons. Accordingly, eighty studies were included in the systematic review: 15 were qualitative whereas 65 were quantitative.

IV. DATA SYNTHESIS

A. Primary Studies Overview

Eighty studies formed the basis of the study: 20 on QAS based on syntax and context; 20 on QAS based on word encoding and knowledge systems; 20 based on forms of deep learning [73, 74; and 20 based on modern components of machine learning and artificial intelligence. It was imperative to select an equal number of studies in each of the four categories to highlight the main directions in QAS research. In addition to being relevant to the research questions, the selected studies adhered to the inclusion and exclusion criteria. This means that the selected studies were authored in English, published after January 1, 2018, and were academic research work found in scholarly journals and conferences. Out of the eighty studies surveyed, only sixty-nine were primary studies.

B. Answering the Research Questions

This section discusses the relationship between the selected studies and the research questions. The relevant research articles extracted are utilized to answer each research question as shown in Table III.

TABLE III.   STUDIES RELEVANCE TO RQS

RQ No.	No. of Studies
RQ1	69
RQ2	55
RQ3	41

This research paper conducted a systematic review of literature, rather than the conventional literature review, due to the need to attain scholarly vigor. In addition to enabling the researcher to obtain the most relevant studies, systematic reviews of literature adopt an objective perspective, which limits biases and enhances the usability of the research findings. Generally, systematic reviews of literature are implemented to summarize existing evidence in a given area, identify gaps for further investigation, and provide a framework for positioning new research activities. In this study, appropriate search terms were utilized in conjunction with various Boolean operators and search strategies to obtain studies to answer the three research questions.

The first research question (RQ1) focused on providing insights concerning the current state of QAS research. The idea is to provide the current understanding of QAS systems in terms of the approaches utilized, their effectiveness and accuracy, and potential areas of improvement. Accordingly, this study explored studies published after January 1, 2018. In total, sixty-nine studies were identified and examined to provide contemporary understanding of QAS research.

The second research question (RQ2) aimed at identifying the most significant gaps and limitations in the reviewed studies. One of the major gaps and limitations is the inability of the developed QAS systems to be utilized for a variety of tasks [1, 4, 6, 7, 15, 24, 30, 74, 75, 80, 81, 82, 83, 84, and 85]. From an ideal standpoint, a QAS system should be applicable to different questions and settings. For example, Utomo [4] developed a QAS system that can only be used with the Quran. Another critical limitation is that a typical QAS system exhibits weaknesses associated with the model or algorithm used [4, 23, 29, 31, 42, 47, 60]. For example, Jovita developed a model that required a long time to give an answer (about 29 seconds) [42]. Besides efficiency, some models are associated with poor precision. Deep learning models, in particular, require quality and large volumes of training data [47]. Thirdly, some of the QAS systems require question templates, selection of hot terms, and standard datasets, which limits their applicability in the practical environment [8, 13, 16, 20, 21, 22, 27, 28, 34, 36, 39, 48, 49, 62, 69]. Other limitations of the studies relate to the thoroughness of the assessments and explanations of the QAS

systems developed [11. 26, 40, 63, 72, 79]. For instance, Abdiansah developed a QAS system that was tested on only three search engines [11].

The final research question (RQ3) aimed at identifying the most effective techniques utilized in the design of QAS systems. Based on the search conducted, the four most effective approaches are syntax and context; word encoding and knowledge systems; forms of deep learning; and components of machine learning and artificial intelligence. Each of the four approaches was studied using 20 research articles. The syntax and context approach places questions within their context, both in terms of the semantic information carried by noun, preposition, and verb phrases and other syntactic entities, as well as the discourse roles relating to the entire question-answering activity. The word encoding and knowledge technique entails utilizing knowledge bases, in combination with question encoding at the character level or using word embedding, to answer a question. The deep learning technique encompasses multiple layers of algorithms to progressively extract high-level features from a question to enable accurate answering. Finally, some question answering systems employ diverse components of machine learning and artificial intelligence, other than deep learning, to answer questions.

V. FINDINGS

This systematic review of literature demonstrates that the current state of QAS research is highly divergent. It appears that different scholars are setting out to develop their individual QAS systems from scratch. This trend could be explained by the fact that QAS is an emerging field. The diversity in QAS techniques means that it is almost impossible to compare them objectively. There is also an emerging trend of combining different components, which makes it difficult to evaluate the effect of each component individually. Accordingly, the adoption of a modular approach could be helpful as it would enable the scientific community to contribute by developing new plugins to improve or replace existing ones. Despite the challenges, QAS research is making positive strides towards the creation of accurate question answering systems.

The review also highlights various significant gaps and limitations in QAS research. A key limitation identified is the highly focused nature of the QAS developed. In addition, the models utilized have weaknesses, which limits the accuracy and efficiency of the entire QAS. Deep learning models, while suited to QAS applications, require vast amounts of quality training data during their development [61, 62, 63, 64, 65, 66, 67, and 68]. Without such data, their effectiveness and applicability reduce significantly. In research studies, particularly those targeting machine learning, the availability of unbiased training data is often a challenge. Moreover, some of the QAS developed only work well with standard datasets. Accordingly, when testing them, researchers are likely to obtain high accuracies. However, in practical settings, standard datasets are unavailable. Furthermore, the research methodologies adopted by the different scholars were deficient as some of the QAS developed were not evaluated comprehensively.

The design of QAS can take one of the four approaches identified. The first one encompasses examining the syntax of the question and the context in which it is placed to enable accurate answering. The second approach involves encoding words contained in the question and then utilizing knowledge bases to find the correct answer. The third method entails utilizing some forms of deep learning, which enables a progressive extraction of information from questions during the answering process. Fourthly, artificial intelligence and machine learning are routinely applied in question answering systems.

VI. DISCUSSION

A. Research Limitations

The findings of this study must be understood within the limitations encountered. One major weakness is that the systematic review of literature was limited to studies published in English. While this requirement was necessary to ensure that the selected studies were understandable to the author, it is possible that some helpful studies were eliminated. Another limitation is that only QAS studies published after January 1, 2018 were surveyed. This criterion might also have limited the inclusion of potentially helpful studies despite being fairly older. Besides, the systematic review of literature included studies that had different methodological weaknesses, including inadequate QAS system evaluation. Accordingly, limitations in individual studies affected the overall strength of this systematic review.

B. Research Conclusion

The objectives of this study entailed providing a picture of the current state of QAS research, discuss gaps and limitations in the QAS research, and explore effective methods utilized in the design of QAS. The study adopted the systematic review of literature research methodology and encompassed examining relevant studies published in English after January 1, 2018. A total of eighty studies were selected for the research study. However, only 69 were relevant to the first research question, 55 were relevant to the second research question, and 41 answered the final research question. Based on the findings, QAS research literature is growing but is highly divergent as scholars adopt different techniques. The main techniques include syntax and context, word encoding and knowledge systems, deep learning [21, 28, 62, 66, 73, 83], and artificial intelligence and machine learning. Some of the significant gaps identified include ineffectiveness and inefficiencies associated with the models adopted, the highly focused nature of QAS systems developed, the reduced practicality of QAS due to the need for standard datasets or question formats, and the inability to test QAS thoroughly. Future research ought to focus on the development of QAS based on modular approaches to enhance collaboration within the scientific community. Future studies should also examine the applicability of some of the developed QAS in practical environments.

REFERENCES

[1] M. Al-Shanaq, K. Nahar. "Aqas: Arabic Question Answering System Based on Svm, Svd, and Lsi." Journal of Theoretical and Applied Information Technology, vol. 97, no. 2, , pp. 681-91, 2019.

[2] M. Anggraeni. "Literation Hearing Impairment (I-Chat Bot): Natural Language Processing (NLP) and Naïve Bayes Method." Journal of

Physics: Conference Series, vol. 1201, no. 1, 2019. https://doi.org/10.1088/1742-6596/1201/1/012057.

[3] A. Asai, K. Hashimoto, H. Hajishirzi, R. Socher, C. Xiong Asai. "Learning to Retrieve Reasoning Paths over Wikipedia Graph for Question Answering". Proceding of ICLR, pp. 1-22, 2020.http://arxiv.org/abs/1911.10470.

[4] A. Azmi, A. Omar, a. Hussain. "Computational and Natural Language Processing Based Studies of Hadith Literature: A Survey." Artificial Intelligence Review, vol. 52, no. 2, Springer, pp. 1369-414, 2019. https://doi.org/10.1007/s10462-019-09692-w.

[5] A. NS, A. UTAMI. "Information Extraction from Web as Knowledge Resources for Indonesian Question Answering System". In Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019) (pp. 419-425), May 2020. https://doi.org/10.2991/aisr.k.200424.064.

[6] A. Asma, P. Zweigenbaum. "MEANS: A medical question-answering system combining NLP techniques and semantic Web technologies". Journal of information processing & management, Volume 51, Issue 5, Pages 570-594, September 2015.

[7] S. Banerjee, S. Naskar, S. Bandyopadhyay, P. Rosso. "Classifier Combination Approach for Question Classification for Bengali Question Answering System." Sadhana - Academy Proceedings in Engineering Sciences, vol. 44, no. 12, Springer India, 2019, doi:10.1007/s12046-019-1224-8.

[8] R. Bakis, D. Connors, P. Dube, P. Kapanipathi, R. Kumar, D. Malioutov, & C. Venkatramani. "Performance of natural language classifiers in a question-answering system" . IBM Journal of Research and Development, 61(4):14:1-14:10, 2017. https://doi.org/10.1147/JRD.2017.2711719.

[9] P. Baudiš, J. Šedivý. "Modeling of the question answering task in the yodaqa system". In International Conference of the Cross-Language Evaluation Forum for European Languages. September 2015. https://doi.org/10.1007/978-3-319-24027-5_20.

[10] A. Bhandwaldar, W. Zadrozny. "UNCC QA: biomedical question answering system". In Proceedings of the 6th BioASQ Workshop A challenge on large-scale biomedical semantic indexing and question answering. November 2018. https://doi.org/10.18653/v1/W18-5308.

[11] C. Soares, M. Antonio, and F. Parreiras. "A on Question Answering Techniques, Paradigms and Systems." Journal of King Saud University - Computer and Information Sciences, vol. 32, no. 6, King Saud University, pp. 635-46, 2020. https://doi.org/10.1016/j.jksuci.2018.08.005.

[12] A. Chen, G. Stanovsky, S. Singh, M. Gardner "Evaluating Question Answering Evaluation". Proceedings of the 2nd Workshop on Machine Reading for Question Answeringpp. 119-24, doi:10.18653/v1/d19-5817, 2019. https://doi.org/10.18653/v1/D19-5817.

[13] C. Chen, C. Wu, C. Lo, F. Hwang. "An augmented reality question answering system based on ensemble neural networks". IEEE Access, August 2017. https://doi.org/10.1109/ACCESS.2017.2743746.

[14] W. Cui, Y. Xiao, H. Wang, Y. Song, S. Hwang, W. Wang. "KBQA: Learning Question Answering over QA Corpora and Knowledge Bases." Proceedings of the VLDB Endowment, vol. 10, no. 5, pp. 565-576, 2016. doi:10.14778/3055540.3055549.

[15] M. Dehghani, H. Azarbonyad, K. Kamps, M. Rijke. "Learning to Transform, Combine, and Reason in Open Domain Question Answering." CEUR Workshop Proceedings, vol. 2491, 2019. https://doi.org/10.1145/3289600.3291012.

[16] D. Fushman, Y. Mrabet, A. Abacha. "Consumer Health Information and Question Answering: Helping Consumers Find Answers to Their Health-Related Information Needs." Journal of the American Medical Informatics Association, vol. 27, no. 2, Oxford University Press, 2020, pp. 194-201. https://doi.org/10.1093/jamia/ocz152.

[17] D. Diefenbach, A. Both, K. Singh, Pierre Maret. "Towards a question answering system over the semantic web". Semantic Web, pp. 1-16, 2018. https://doi.org/10.3233/SW-190343.

[18] E. Dimitrakis, K. Sgontzos, Y. Tzitzikas. "A Survey on Question Answering Systems over Linked Data and Documents." Journal of Intelligent Information Systems, vol. 55, no. 2, pp. 233-59, 2020. doi:10.1007/s10844-019-00584-7.

[19] T. Dodiya, S. Jain. "Question classification for medical domain question answering system". In 2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), December 2016. https://doi.org/10.1109/WIECON-ECE.2016.8009118.

[20] M. Espositoa, E. Damianoa, A. Minutoloa, G. De Pietroa, H. Fujitaal. "Hybrid Query Expansion Using Lexical Resources and Word Embeddings for Sentence Retrieval in Question Answering." Information Sciences, vol. 514, Elsevier, 2020, pp. 88-105. https://doi.org/10.1016/j.ins.2019.12.002.

[21] S. Gupta, N. Khade. "BERT Based Multilingual Machine Comprehension in English and Hindi". Special Issue on Deep Learning of ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP) no. 1, pp. 1-13, 2020.

[22] H. Alami, N. Ennahnahi, K. Zidani. "An Arabic Question Classification Method Based on New Taxonomy and Continuous Distributed Representation of Words." Journal of King Saud University-Computer and Information Sciences, Elsevier, 2019.

[23] S. Hamed, M. Ab Aziz. "A Question Answering System on Holy Quran Translation Based on Question Expansion Technique and Neural Network Classification." Journal of Computer Science, vol 12(3):169-177, January 2016. https://doi.org/10.3844/jcssp.2016.169.177.

[24] S. Hazrina, N. Sharef, H. Ibrahim, M. AzmiMurad, S. MohdNoah. " Review on the advancements of disambiguation in semantic question answering system". Information Processing & Management, 2017. https://doi.org/10.1016/j.ipm.2016.06.006.

[25] H. Xiao, X. Huang, J. Zhang, D. Li, P. Li. "Knowledge Graph Embedding Based Question Answering." WSDM 2019 - Proceedings of the 12th ACM International Conference on Web Search and Data Mining, no. Ccl, pp. 105-13, 2019. doi:10.1145/3289600.3290956.

[26] D. Hudson, C. Manning. "GQA: A New Dataset for Real-World Visual Reasoning and Compositional Question Answering." Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. June 2019, pp. 6693-702, 2019. doi:10.1109/CVPR.2019.00686.

[27] S. Jaya, M. Zidny , M. Gunawan. "Integrated Passage Retrieval with Fuzzy Logic for Indonesian Question Answering System." International Journal on Perceptive and Cognitive Computing, vol. 5, no. 2, pp. 31-34, 2019. doi:10.31436/ijpcc.v5i2.118.

[28] K. Karpagam, K. Madusudanan, A. Saradha. "Deep Learning Approaches for Answer Selection in Question Answering System for Conversation Agents." ICTACT Journal on Soft Computing, vol. 10, no. 2, pp. 2040-44, 2020. doi:10.21917/ijsc.2020.0289.

[29] Kitchenham, B. Kitchenham, O. Brereton, D. Budgen, M. Turner, J. Bailey, S. Linkman. "Systematic Literature Reviews in Software Engineering - A Systematic Literature Review." Information and Software Technology, vol. 51, no. 1, pp. 7-15, 2009. doi:10.1016/j.infsof.2008.09.009.

[30] L. Kodra, K. Elinda. "Question Answering Systems: A Review on Present Developments, Challenges and Trends." International Journal of Advanced Computer Science and Applications, vol. 8, no. 9, pp. 217-24, 2017. doi:10.14569/ijacsa.2017.080931.

[31] M. Kowsher, M. Rahman, S. Ahmed, N. Prottasha. "Bangla Intelligence Question Answering System Based on Mathematics and Statistics." 2019 22nd International Conference on Computer and Information Technology (ICCIT), pp. 1-6, 2019. doi:10.1109/ICCIT48885.2019.9038332.

[32] K. Krishna, M. Iyyer. "Generating Question-Answer Hierarchies." ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference, pp. 2321-34, 2020. doi:10.18653/v1/p19-1224.

[33] P. Rosso, Y. Benajiba, A. Lyhyaoui. "Towards a Passages Extraction Method for Arabic Question Answering Systems." International Conference on Advanced Intelligent Systems for Sustainable Development, Springer, pp. 230-37, 2019. https://doi.org/10.1007/978-3-030-36653-7_23.

[34] G. Popek, W. Lorkiewicz. "Grounding of Modal Responses in Question Answering System Equipped with Hierarchical Categorisation." Procedia Computer Science, vol. 176, Elsevier B.V., pp. 3163-72, 2020. doi:10.1016/j.procs.2020.09.172.

[35] T. Abedissa, M. Libsie. "Amharic Question Answering for Biography, Definition, and Description Questions". Information and Communication Technology for Development for Africa, vol 1026. Springer, Cham, 2019. doi.org/10.1007/978-3-030-26630-1_26.

[36] B. McCann, N. Keskar, C. Xiong, R. Socher. "The Natural Language Decathlon: Multitask Learning as Question Answering". 2018, http://arxiv.org/abs/1806.08730.

[37] H. Mozannar, E. Maamary, K. El Hajal, H. Hajj. "Neural Arabic Question Answering". Proceedings of the Fourth Arabic Natural Language Processing Workshop, pp. 108-18, August 2019. doi:10.18653/v1/w19-4612.

[38] G. Nanda, M. Dua and K. Singla, "A Hindi Question Answering System using Machine Learning approach". International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT), New Delhi, India, pp. 311-314, 2016. doi: 10.1109/ICCTICT.2016.7514599.

[39] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh and K. Shaalan, "Speech Recognition Using Deep Neural Networks: A Systematic Review," in *IEEE Access*, vol. 7, pp. 19143-19165, 2019. doi: 10.1109/ACCESS.2019.2896880.

[40] O.Bolanle, and E.Adebisi, "A Review of Question Answering Systems," Journal of Web Engineering, vol. 17, no. 8, pp. 717-58,2019.

[41] O.Chitu, and K.Schabram, "A Guide to Conducting a Systematic of Information Systems Research," SSRN Electronic Journal, vol. 10, pp. 1-3, 2012.

[42] P. V. Rajaraman, and M.Prakash, "A Survey on Text Question Responsive Systems in English and Indian Languages," International Conf. on Soft Computing and Signal Processing, Springer, pp. 267-77, 2019.

[43] P.Ranjan, and R. C. Balabantaray, " Question answering system for factoid-based question," 2nd International Conf. on Contemporary Computing and Informatics (IC3I), IEEE, 2016.

[44] S.K.Ray, A. Ahmad, and K. Shaalan, "A Review of the State of the Art in Hindi Question Answering Systems," Intelligent Natural Language Processing: Trends and Applications, Springer, pp. 265-92, 2018.

[45] A.Salunkhe, "Evolution of Techniques for Question Answering over Knowledge Base: A Survey," International Journal of Computer Applications, vol. 177, no. 34, pp. 9-14, 2020.

[46] V.Sharma, N.Kulkarni, S.Pranavi, G.Bayomi, E.Nyberg and T. Mitamura, "BioAMA: Towards an End to End BioMedical Question Answering System," In Proc. of the BioNLP 2018 workshop, July ,2018.

[47] Si.Shijing, W.Zheng , L.Zhou, and M.Zhang, "Sentence Similarity Computation in Question Answering Robot," Journal of Physics: Conf. Series, vol. 1237, no. 2, pp 320-355, 2019.

[48] K.Singh, S.A. Radhakrishna, and A. Both, "Why Reinvent the Wheel: Let's Build Question Answering Systems Together," Proce. of the World Wide Web Conf., The Web Conf., pp. 1247-56, 2018.

[49] K.Singh, A.Both, D.Diefenbach, S.Shekarpour, D.Cherix, and C. Lange, "Qanary-the fast track to creating a question answering system with linked data technology," In European Semantic Web Conf., May, 2016.

[50] R.A. Stein, P.A.Jaques, and J. F. Valiati, "An Analysis of Hierarchical Text Classification Using Word Embeddings," Information Sciences, vol. 471, pp. 216-32, 2019.

[51] I.Thalib, , and S. Indah, "A Review on Question Analysis, Document Retrieval and Answer Extraction Method in Question Answering System," International Conf. on Smart Technology and Applications (ICoSTA), IEEE, pp. 1-5, 2020.

[52] N.Tohidi, and S.Hossain, "Multi-Objective Question Answering System," Journal of Intelligent and Fuzzy Systems, vol.36, no.4, pp.3495-3512, 2019.

[53] F.S.Utomo, N. Suryana, and N.Azmi, "New Instances Classification Framework on Quran Ontology Applied to Question Answering System," Telkomnika (Telecommunication Computing Electronics and Control), Indonesian Journal of Electrical Engineering, vol. 17, no. 1, pp. 139-46, 2019.

[54] B.Wanjawa, and L.Muchemi, "Question Answering Using Automatically Generated Semantic Networks-the Case of Swahili Questions," IST-Africa Conference (IST-Africa), IEEE, pp. 1-8, 2020.

[55] J. Yu, M.Qiu, J.Jiang, J. Huang, and S. Song, "Modelling Domain Relationships for Transfer Learning on Retrieval-Based Question Answering Systems in E-Commerce," Proc. of the 11th ACM International Conf. on Web Search and Data Mining, WSDM, Febuary, no. 1, pp. 682-90, 2018.

[56] D.Savenkov, and E.Agichtein. "Crowd-powered real-time automatic question answering system," Fourth AAAI Conf. on Human Computation and Crowdsourcing, pp. 189-199, 2016.

[57] Z.Abbasiyantaeb, and S. Momtazi, "Text-based question answering from information retrieval and deep neural network perspectives: A survey," Advanced Review, 2020.

[58] P.Baudiš, and J. Šedivý, "Modeling of the question answering task in the YodaQA system," International Conf. of the Cross-Language Evaluation Forum for European Languages, Springer, Cham, pp. 222-228, 2015.

[59] A.Bouziane,D.Bouchiha,N.Doumi, and M.Malki, "Question answering systems: Survey and trends," Procedia Computer Science, vol. 73, pp. 366-375, 2015.

[60] V.Datla, S. A. Hasan, Joey Liu, Y. Benajiba, K.Lee, A.Qadir, A.Prakash , and O.Farri, "Open Domain Real-Time Question Answering Based on Semantic and Syntactic Question Similarity," Journal of Web Engineering (JWE), vol.17, no.8, pp.717-758, 2016.

[61] K.Höffner, S.Walter, E.Marx, J.Lehmann, A.N.Ngomo, and R.Usbeck, "Overcoming challenges of semantic question answering in the semantic web," Semantic Web Journal, pp. 1-21, 2016.

[62] L.Tuan, T.Bui, and S.Li, "A review on deep learning techniques applied to answer selection," Proc. of the 27th international Conf. on computational linguistics, pp. 2132-2144, 2018.

[63] L.Juan, C.Zhang, and Z.Niu, "Answer extraction based on merging score strategy of hot terms," Chinese Journal of Electronics, vol. 25, no. 4, pp. 614-620, 2016.

[64] M.Amit, and S.K.Jain, "A survey on question answering systems with classification," Journal of King Saud University-Computer and Information Sciences, vol. 28, no. 3, pp. 345-361, 2016.

[65] M.Ajitkumar, S. Khillare, and C. Namrata, "Question answering system, approaches and techniques: A review," International Journal of Computer Applications, vol. 141, no. 3, pp. 0975-8887, 2016.

[66] S.Yashvardhan, and S.Gupta, "Deep learning approaches for question answering system," Procedia Computer Science, vol. 132, pp. 785-794, 2018.

[67] S.Anjali, and P. K. Yadav, "A survey on question-answering system," International Journal of Engineering and Computer Science, vol. 6, no. 3, pp. 345-361, 2017.

[68] S.Anbuselvan, M.Muniandy, and L.E.Heng, "Question Classification Using Statistical Approach: A Complete Review," Journal of Theoretical & Applied Information Technology, vol. 71, no. 3, pp 386-395, 2015.

[69] M.Schubotz, P.Scharpf, K.Dudhat, Y.Nagar, F.Hamborg, and B.Gipp, , "Introducing A Math-Aware Question Answering System," Journal of Information Discovery and Delivery, vol.46, issue 4, pp. 214-224, 2018.

[70] F.Schulze, R.Schüler, T.Draeger, D.Dummer, A.Ernst, P.Flemming, M.Neves, "Hpi question answering system in bioasq," In Proceedings of the Fourth BioASQ workshop, pp 38-44, August, 2016.

[71] S.Shekarpour, E.Marx, A.C.N.Ngomo, and S.Aue, "Sina: Semantic interpretation of user queries for question answering on interlinked data," Journal of Web Semantics vol. 30, pp. 39-51, 2015.

[72] D.Su, Y.Xu, G. I.Winata, P.Xu, H.Kim, Z.Liu, and P.Fung, "Generalizing Question Answering System with Pre-trained Language Model Fine-tuning," In Proceedings of the 2nd Workshop on Machine Reading for Question Answering, November ,2019.

[73] M.Toshevska, G.Mirceva, and M.Jovanov, "Question Answering with Deep Learning: A Survey,"16th International Conf. on Informatics and Information Technologies, CIIT, 2019.

[74] X.Yang, and P.Liu, "Question recommendation and answer extraction in question answering community," International Journal of Database Theory and Application, vol. 9, no. 1, pp. 35-44, 2016.

[75] W.Yang, Y.Xie, A. Lin, X.Li, L.Tan, K.Xiong, and J. Lin, "End-to-end open-domain question answering with bertserini," Proceedings of the

2019 Conf.of the North American Chapter of the Association for Computational Linguistics (Demonstrations), pp.v2, 2019.

[76] Y.Yang, W. T.Yih, and C.Meek, "Wikiqa: A challenge dataset for open-domain question answering," In Proceedings of the 2015 Conf. on empirical methods in natural language processing, Association for Computational Linguistics, pp.2013-2018, 2015.

[77] Y. T.Yeh, and Y. N. Chen, " Qainfomax: Learning robust question answering system by mutual information maximization," Computer Science, Mathematics, 2019.

[78] Y.Deepa, T. N. Manjunath, and R.S. Hegadi, "A Survey of Intelligent Question Answering System Using NLP and Information Retrieval Techniques," International Journal of Advanced Research in Computer and Communication Engineering vol. 5, no. 5, pp. 536-540,2016.

[79] W.Yu, L.Wu, Y.Deng, R.Mahindru, Q.Zeng, S.Guven, and M.Jiang, "A Technical Question Answering System with Transfer Learning," In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, System Demonstrations, October, 2020.

[80] A. Clementeena, and P. Sripriya, "A literature survey on question answering system in Natural Language Processing," International Journal of Engineering and Technology, vol. 7, no. 2, 2018, pp. 452-455.

[81] A.Ansari, M.Maknojia, and A.Shaikh, "Intelligent question answering system based on artificial neural network," IEEE International Conf. on Engineering and Technology, IEEE, Coimbatore, India,2016.

[82] L.Jovita, A.Hartawan, and D.Suhartono, "Using vector space model in question answering system," Procedia Computer Science vol. 59, pp. 305 - 311, 2015. https://doi.org/10.1016/j.procs.2015.07.570.

[83] Z. Huang et al., "Recent Trends in Deep Learning Based Open-Domain Textual Question Answering Systems," in IEEE Access, vol. 8, pp. 94341-94356, 2020. doi: 10.1109/ACCESS.2020.2988903.

[84] R. Chakravarti, A. Ferritto, B. Iyer, L. Pan, R. Florian, S. Roukos and A. Sil. "Towards building a Robust Industry-scale Question Answering System." Processsding of the 28th international conference on computational ligusitics: industry track, p.p. 90-101, 2020.

[85] BB Cambazoglu, M Sanderson, F Scholer, B Croft - sigir.org. " A Review of Public Datasets in Question Answering Research". ACM SIGIR Forum. Vol. 54 No. 2. December 2020.